Supporting Information

Targeted Capture and Heterologous Expression of the *Pseudoalteromonas* Alterochromide Gene Cluster in *Escherichia coli* Represents a Promising Natural Product Exploratory Platform

Avena C. Ross, Lauren E. S. Gulland, Pieter C. Dorrestein, Bradley S. Moore

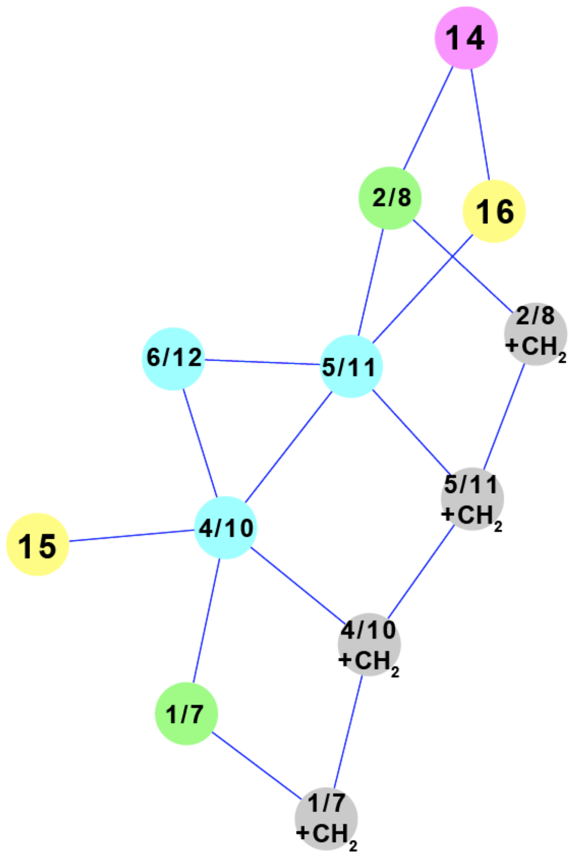
Table of Contents

Table A: antiSMASH[1] analysis of *Pseudoalteromonas* genome sequence data to identify types of natural product pathways present

Pseudoalteromonas strain (genome sequence-number of scaffolds)	RiPPs	Non-Ribosomal Peptide (NRP)	Polyketide (PK)	Hybrid NRP/PK	Homo serine lactone	Indole	Siderophore	Other
P. agrivorans S816 (133 scaffolds)	1						1	
P. arctica A3712 (152 scaffolds)	1							
P. atlantica T6c (1 scaffold)	1				2			
P. atlantica TB41 (122 scaffolds)	1			1			1	
P. citrea NCIMB 1889 (223 scaffolds)	5	4	2	4				3
P. citrea (314 scaffolds)	3	1						1
P. flavipulchra JG1 (61 scaffolds)	5	13		3				
P. haloplanktis (11 strains on 2- 565 scaffolds*)*	21						1	
P. lipolytica SCSIO 04301 (12 scaffolds)	1							
P. luteoviolacea 2ta16 (137 scaffolds)	3	3	2	5	1	1		1
P. luteoviolacea B ATCC 29581 (61 scaffolds)	3	1				1		1
P. marina mano4 (69 scaffolds)	3							
P. piscicida ATCC 15057 (58 scaffolds)	4	3		4				
P. piscicida JCM 20779 (207 scaffolds)	4	4	4					
P. rubra ATCC 29570 (194 scaffolds)	3	17	1	4	1			3
P. ruthenica CP76 (120 scaffolds)	2						1	
P. spongiae UST010723-006 (56	1							

scaffolds)								
P. tunicata D2 (37 scaffolds)	1	2				1	1	1
P. undina NCIMB 2128 (49 scaffolds)	1						1	
Pseudoalteromonas sp. (18 strains on 2-369 scaffolds)	33	18		4			9	3

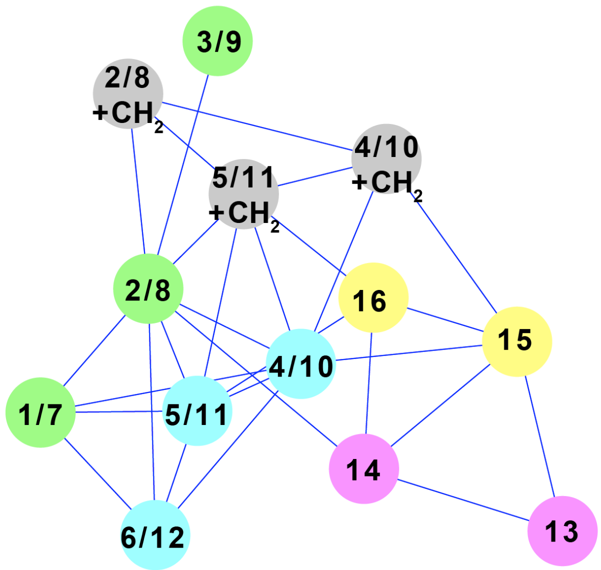
Note- Whole genome shotgun sequencing data downloaded from http://www.ncbi.nlm.nih.gov/ was submitted for analysis by antiSMASH, pathways spread over several scaffolds may be annotated as separate pathways resulting in some over estimation of pathway numbers.



Node Label	Mass m/z	Samples Mass Detected in
1/7	766.379	1,3
2/8	844.329	1,2,3
4/10	792.395	1,3,4
5/11	870.449	1,3
6/12	950.215	1
14	830.274	1
15	778.381	1
16	856.286	1
1/7 + CH_2	780.398	1,3
2/8 + CH_2	858.488	1,3
4/10 + CH_2	806.41	1,3
5/11 + CH_2	884.484	1,3

Sample 1: *P. piscicida*
Sample 2: *E. coli* pACR02/pACR10
Sample 3: *E. coli* pACR07
Sample 4: *E. coli* pACR02
Sample 5: *E. coli* pEtDuet-1

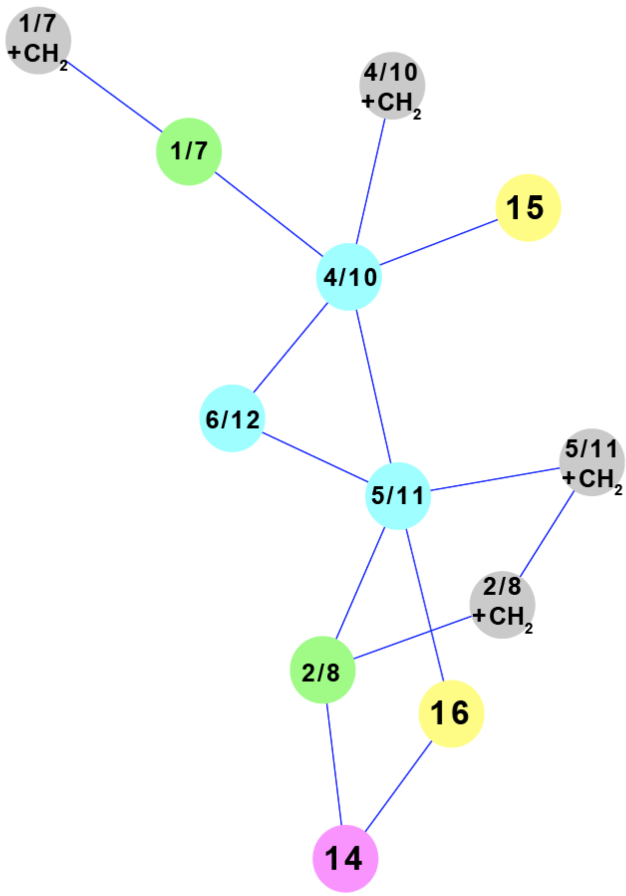
Figure S1: Mass spectral molecular network cluster showing alterochromide molecules produced by *Pseudoalteromonas piscicida* and *Escherichia coli* heterologous expression using pACR02 alone, pACR02/pACR10 or pACR07. Green nodes represent the A and A′ molecules, blue nodes represent the B and B′ molecules, the pink node represents an A′′ molecule, the yellow nodes represent the B′′ molecules and the grey nodes represent analogues with an additional methylene group at an undetermined location (See Figure S4 for chemical structures of 1-16). Procedure- Mass spectral data was converted to .mzXML format and data was analyzed using a molecular networking workflow described previously,[2] with the following settings: Parent Mass Tolerance: 2.0 Da, Ion Tolerance: 0.5 Da, Min Pairs Cos: 0.7, Min Matched Peaks: 6, Network TopK: 10, Minimum Cluster Size: 2, Maximum Connected Component Size: 0. The molecular network was visualized using un-weighted force directed layout within Cytoscape version 2.8.2.[3]



Node Label	Mass m/z	Samples Mass Detected in
1/7	766.377	1,5
2/8	844.29	1,2,3,6
3/9	924.197	3
4/10	792.393	1, 5
5/11	870.302	1,3
6/12	950.211	1
13	752.362	1
14	830.271	1
15	778.376	1
16	856.287	1
2/8 + CH_2	858.302	1,3
4/10 + CH_2	806.408	1,3,5
5/11 + CH_2	884.317	1

Sample 1: *P. piscicida*
Sample 2: *E. coli* pACR02/pACR10
Sample 3: *E. coli* pACR07
Sample 4: *E. coli* pACR08 (Δ*altA*)
Sample 5: *E. coli* pACR09 (Δ*altN*)
Sample 6: *E. coli* pACR08 (Δ*altA*) + coumaric acid

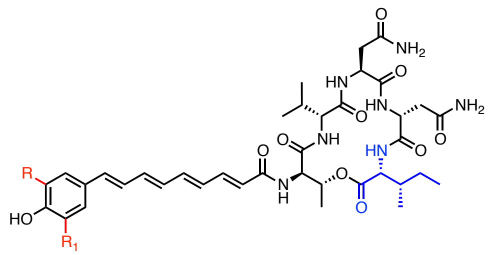
Figure S2: Mass spectral molecular network cluster showing alterochromide molecules produced by *Pseudoalteromonas piscicida*, *Escherichia coli* heterologous expression using pACR02/pACR10 or pACR07 and pathway mutant experiments. Green nodes represent the A and A′ molecules, blue nodes represent the B and B′ molecules, the pink nodes represents A′′ molecules, the yellow nodes represent the B′′ molecules and the grey nodes represent analogues with an additional methylene group at an undetermined location (See Figure S4 for chemical structures of 1-16). Procedure as described in figure S1.



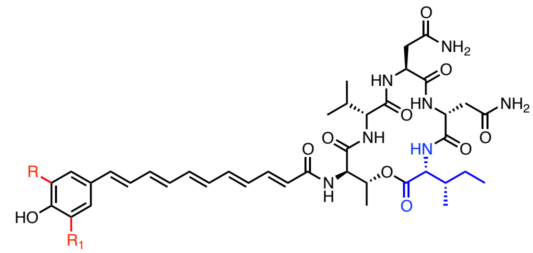
Node Label	Mass m/z	Samples Mass Detected in
1/7	766.379	1,2,3,5
2/8	844.33	1,6
4/10	792.395	1,2,3,5
5/11	870.437	1,2
6/12	950.215	1
14	830.274	1
15	778.377	1,2
16	856.286	1
1/7 + CH_2	780.07	1,2,5
2/8 + CH_2	858.498	1
4/10 + CH_2	806.332	1,2,3,5
5/11 + CH_2	884.593	1,2

Sample 1: *P. piscicida* and *E. coli* pACR07 + Br
Sample 2: *E. coli* pACR07 - Br
Sample 3: *E. coli* pACR09 (ΔaltN) + Br
Sample 4: *E. coli* pACR08 (ΔaltA) + Br
Sample 5: *E. coli* pACR08 (ΔaltA) + coumaric acid - Br
Sample 6: *E. coli* pACR08 (ΔaltA) + coumaric acid + Br

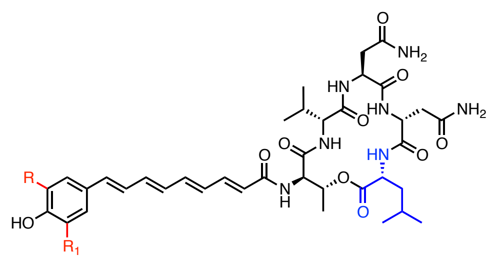
Figure S3: Mass spectral molecular network cluster showing alterochromide molecules produced by *Pseudoalteromonas piscicida*, *Escherichia coli* heterologous expression using pACR02/pACR10 or pACR07 and pathway mutant experiments. Green nodes represent the A and A′ molecules, blue nodes represent the B and B′ molecules, the pink node represents an A′′ molecule, the yellow nodes represent the B′′ molecules and the grey nodes represent analogues with an additional methylene group at an undetermined location (See Figure S4 for chemical structures of 1-16). Procedure as described in figure S1. Note- sample 1 represents extracts of bacteria grown with supplemental bromide, samples 2 and 5 represent extracts of bacteria grown without supplemental bromide.



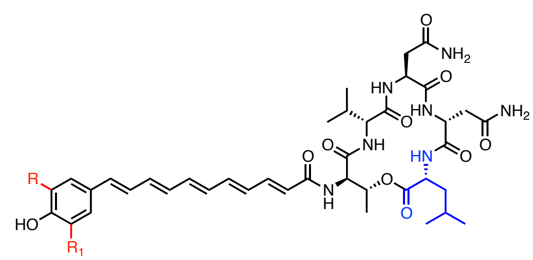
alterochromide A (**1**) R=R_1=H,
Calc. $(M+H)^+$= 766.3776, Obs. $(M+H)^+$= 766.3774
bromoalterochromide A (**2**) R=H, R_1=Br,
Calc. $(M+H)^+$= 844.2881/846.2869, Obs. $(M+H)^+$= 844.2881/846.2863
dibromoalterochromide A (**3**) R=R_1=Br,
Calc. $(M+H)^+$= 924.197, Obs. $(M+H)^+$= 924.1971



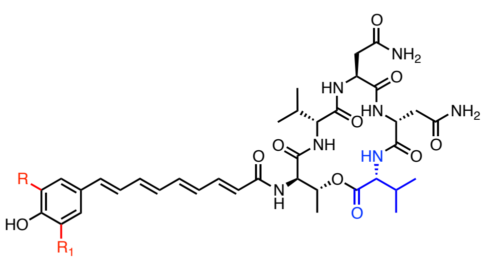
alterochromide B (**4**) R=R_1=H,
Calc. $(M+H)^+$= 792.3932, Obs. $(M+H)^+$= 792.3933
bromoalterochromide B (**5**) R=H, R_1=Br,
Calc. $(M+H)^+$= 870.3037/872.3026, Obs. $(M+H)^+$= 870.3035/872.3027
dibromoalterochromide B (**6**) R=R_1=Br,
Calc. $(M+H)^+$= 950.2127, Obs. $(M+H)^+$= 950.2124



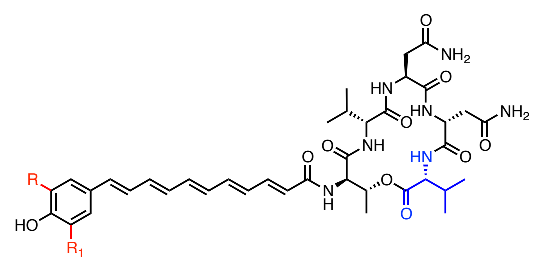
alterochromide A′ (**7**) R=R_1=H,
Calc. $(M+H)^+$= 766.3776, Obs. $(M+H)^+$= 766.3774
bromoalterochromide A′ (**8**) R=H, R_1=Br,
Calc. $(M+H)^+$= 844.2881/846.2869, Obs. $(M+H)^+$= 844.2881/846.2863
dibromoalterochromide A′ (**9**) R=R_1=Br,
Calc. $(M+H)^+$= 924.197, Obs. $(M+H)^+$= 924.1971



alterochromide B′ (**10**) R=R_1=H,
Calc. $(M+H)^+$= 792.3932, Obs. $(M+H)^+$= 792.3933
bromoalterochromide B′ (**11**) R=H, R_1=Br,
Calc. $(M+H)^+$= 870.3037/872.3026, Obs. $(M+H)^+$= 870.3035/872.3027
dibromoalterochromide B′ (**12**) R=R_1=Br,
Calc. $(M+H)^+$= 950.2127, Obs. $(M+H)^+$= 950.2124



alterochromide A″ (**13**) R=R_1=H,
Calc. $(M+H)^+$= 752.3619, Obs. $(M+H)^+$= 752.3617
bromoalterochromide A″ (**14**) R=H, R_1=Br,
Calc. $(M+H)^+$= 830.2724/832.2712, Obs. $(M+H)^+$= 830.2724/832.2714



alterochromide B″ (**15**) R=R_1=H,
Calc. $(M+H)^+$= 778.3776, Obs. $(M+H)^+$= 778.3776
bromoalterochromide B″ (**16**) R=H, R_1=Br,
Calc. $(M+H)^+$= 856.2881/858.2869 Obs. $(M+H)^+$= 856.2886/858.2862

Figure S4: Chemical structure and mass spectral data for alterochromide molecules 1-16. The A and B molecules differ by length of the lipid chain, A/B molecules have an isoleucine as the C-terminal residue, A′/B′ molecules have a leucine as the C-terminal residue and A′′/B′′ molecules have a valine as the C-terminal residue. Stereochemistry of the amino acids in 2/8, 4/10 and 5/11 were determined previously by Kalinovskaya et al. [4] and are in agreement with epimerization domains seen within the gene cluster, by extension stereochemistry for the remaining molecules is assumed to follow the same pattern.

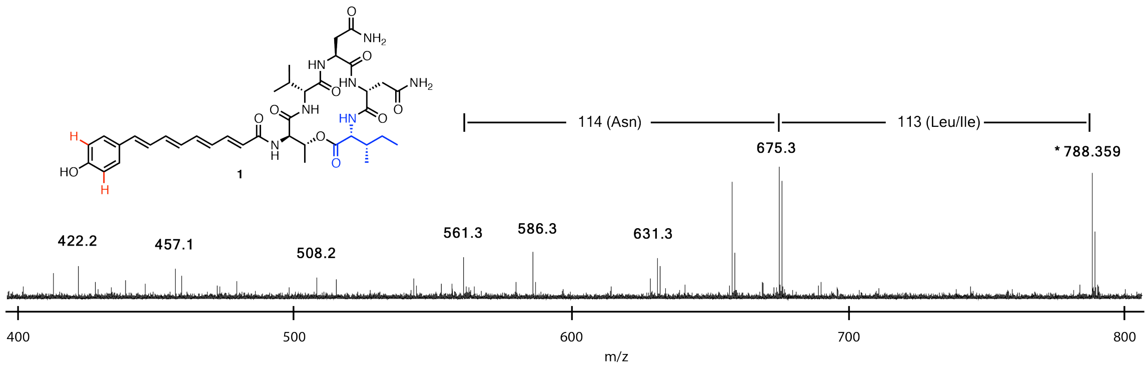


Figure S5: Mass spectral fragmentation data for sodiated adducts of alterochromide A (1) and A′ (7) showing alterochromide Ile/Leu-Asn sequence tag

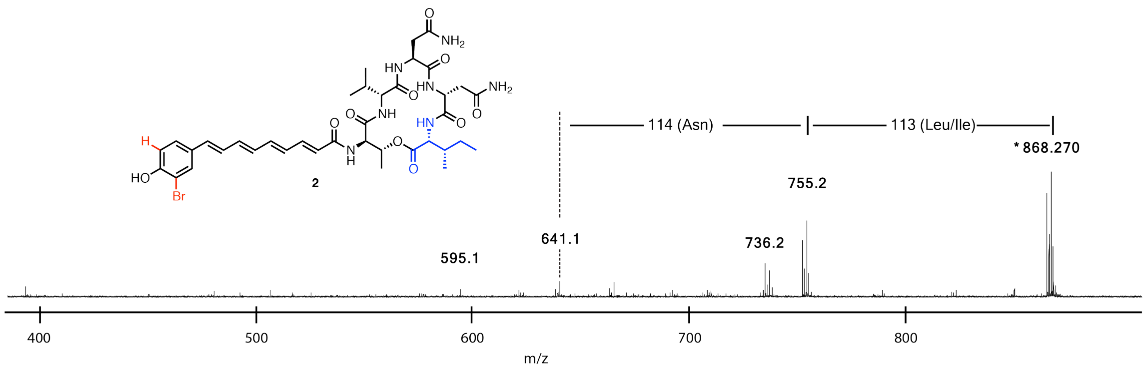


Figure S6: Mass spectral fragmentation data for sodiated adducts of bromoalterochromide A (2) and A′ (8) showing alterochromide Ile/Leu-Asn sequence tag

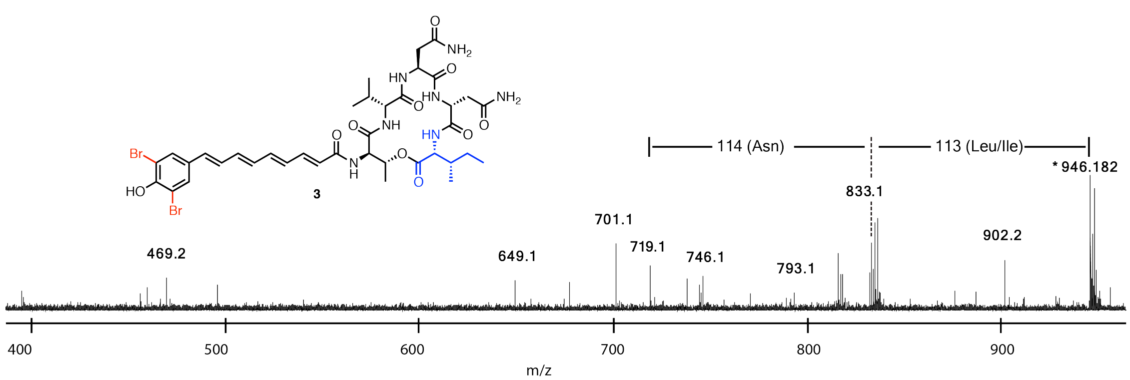


Figure S7: Mass spectral fragmentation data for sodiated adducts of dibromoalterochromide A (3) and A′ (9) showing alterochromide Ile/Leu-Asn sequence tag

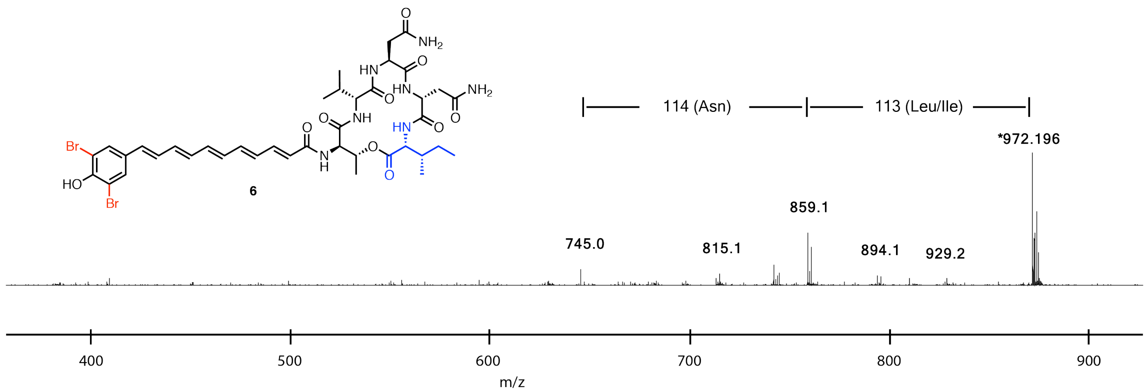


Figure S8: Mass spectral fragmentation data for sodiated adducts of dibromoalterochromide B (6) and B′ (12) showing alterochromide Ile/Leu-Asn sequence tag

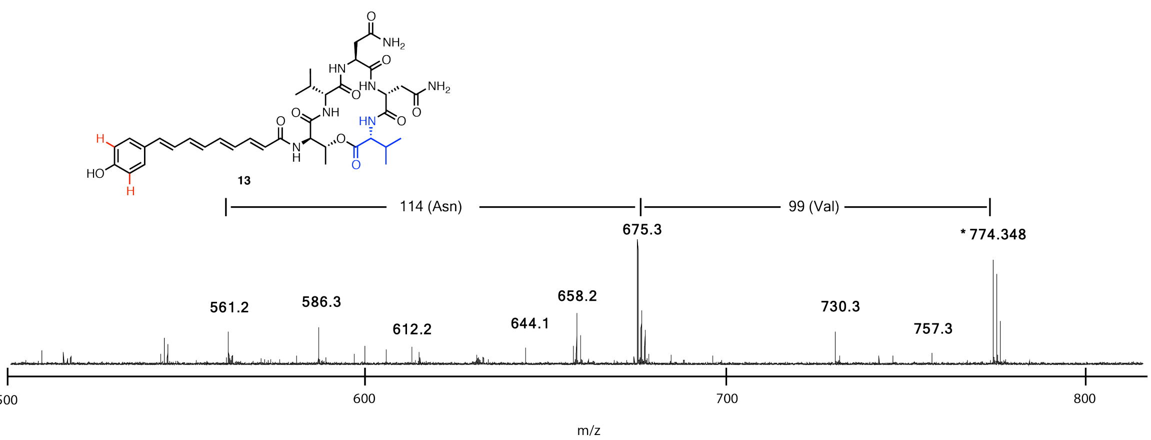


Figure S9: Mass spectral fragmentation data for sodiated adduct of alterochromide A′′ (13) showing alterochromide sequence tag with valine for leucine substitution

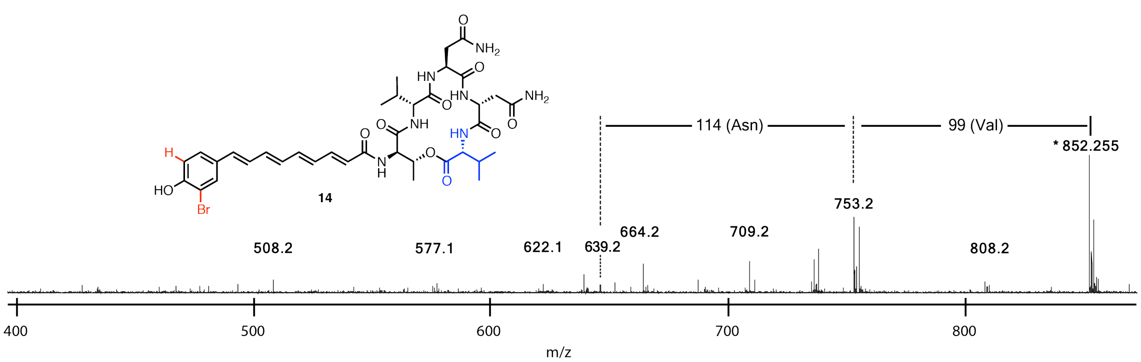


Figure S10: Mass spectral fragmentation data for sodiated adduct of bromoalterochromide A′′ (14) showing alterochromide sequence tag with valine for leucine substitution

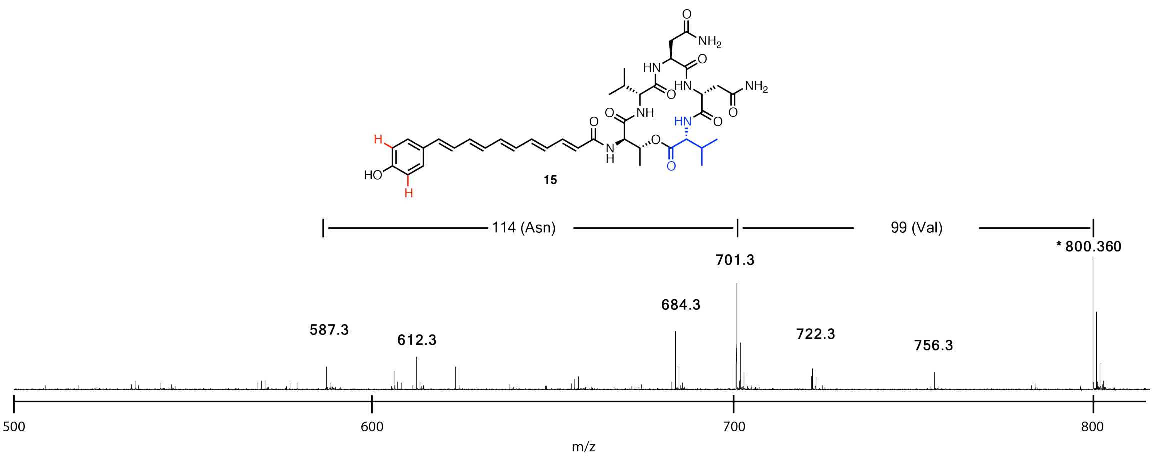


Figure S11: Mass spectral fragmentation data for sodiated adduct of alterochromide B′′ (15) showing alterochromide sequence tag with valine for leucine substitution

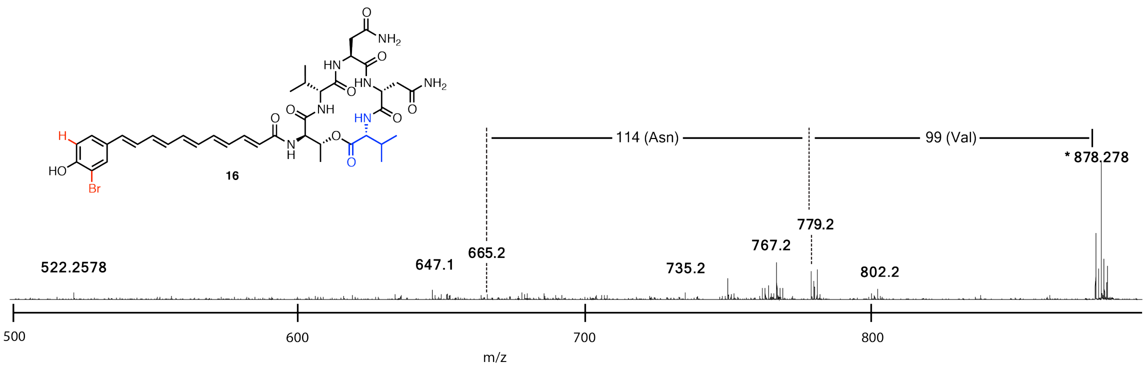


Figure S12: Mass spectral fragmentation data for sodiated adduct of bromoalterochromide B′′ (16) showing alterochromide sequence tag with valine for leucine substitution

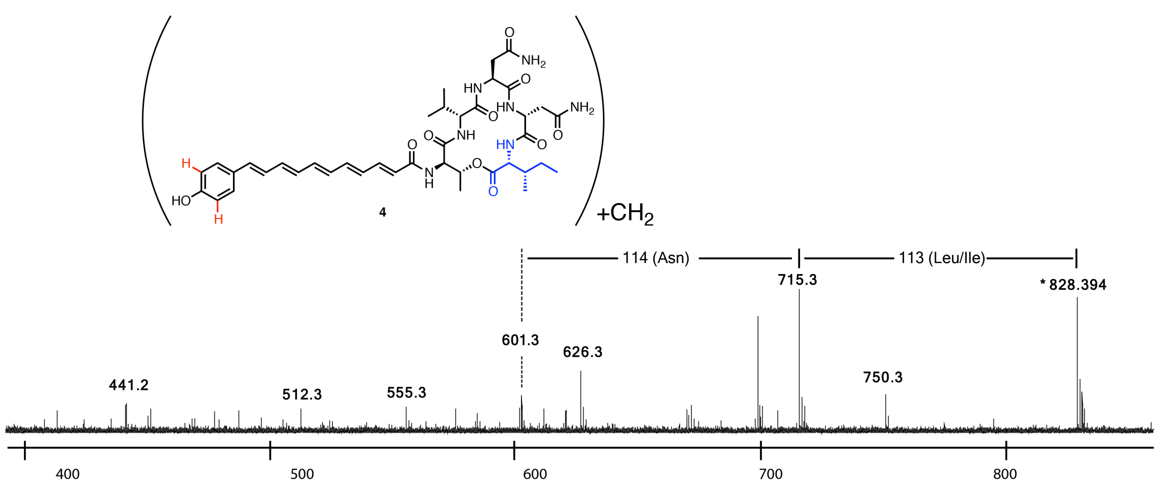


Figure S13: Mass spectral fragmentation data for sodiated adduct of a methylene alterochromide B (4) analogue showing alterochromide Ile/Leu-Asn sequence tag. Site of methylene group is undetermined at present.

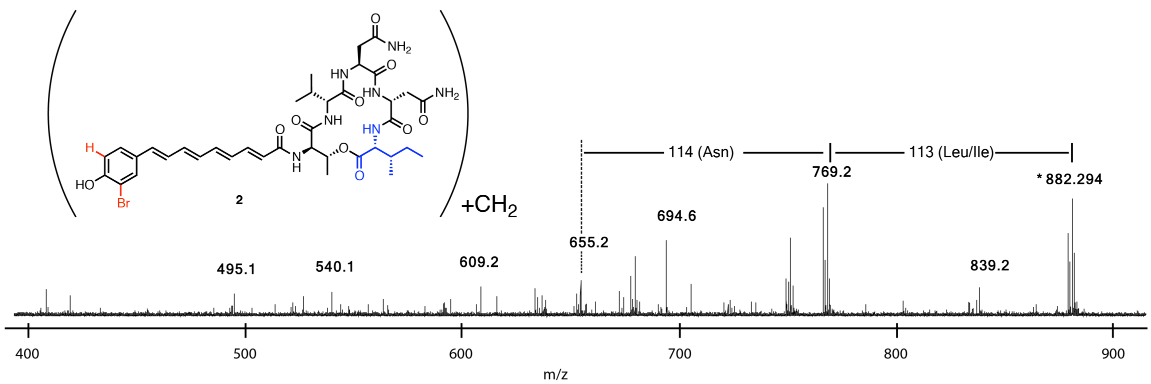


Figure S14: Mass spectral fragmentation data for sodiated adduct of a methylene bromoalterochromide A (2) analogue showing alterochromide Ile/Leu-Asn sequence tag. Site of additional methylene group is undetermined at present.

References

1. **Blin, K., Medema, M. H., Kazempour, D., Fischbach, M. A., Breitling, R., Takano, E., and Weber, T. (2013) antiSMASH 2.0-a versatile platform for genome mining of secondary metabolite producers, *Nucleic Acids Res. 41*, W204-W212.**
2. **Watrous, J. D., Roach, P., Alexandrov, T., Heath, B. S., Yang, J. Y., Kersten, R. D., van der Voort, M., Pogliano, K., Gross, H., Raaijmakers, J. M., Moore, B. S., Laskin, J., Bandeira, N., and Dorrestein, P. C. (2012) Mass spectral molecular networking of living microbial colonies, *Proc. Natl. Acad. Sci. U. S. A. 109*, 1743-1752.**
3. **Smoot, M., Ono, K., Ruscheinski, J., Wang, P.-L., and Ideker, T. (2011) Cytoscape 2.8: new features for data integration and network visualization, *Bioinformatics 27*, 431-432.**
4. **Kalinovskaya, N. I., Dmitrenok, A. S., Kuznetsova, T. A., Frolova, G. M., Christen, R., Laatsch, H., Alexeeva, Y. V., and Ivanova, E. P. (2008) "*Pseudoalteromonas januaria*" SUT 11 as the source of rare lipodepsipeptides *Curr. Microbiol. 56*, 199-207.**